Exhibit B

```
#!/u/bmaison/bin/perl5

open(f,"/u/mangu/C4.5/R8/MyData/two_new_new.data") || die "I can't
";

$maxClustNo=0;
$maxDurNo=0;
$maxnoCand=0;
$maxClustLen=0;

$count=0;

open(fout,">> rules");



while(<f>){
  chop;
  @a=split(/,/);
  $truth[$count]=$a[14];
  $prediction[$count]=1;

  $word1{$a[1]}=1;
  $word2{$a[2]}=1;
  $pair{$a[1]."_".$a[2]}++;
  $epspair{$a[3]."_".$a[4]}=1;

  if ($a[0] > $maxClustNo){
    $maxClustNo=$a[0];
  }
  if ($a[5] > $maxDurNo){
    $maxDurNo=$a[5];
  }
  if ($a[6] > $maxDurNo){
    $maxDurNo=$a[6];
  }

  if ($a[10] > $maxCandNo){
    $maxCandNo = $a[10];
  }
  if ($a[12] > $maxClustLen){
    $maxClustLen = $a[12];
  }
  for ($i=0; $i<=$#a; $i++){
    $sent[$count][$i] = $a[$i];
  }
  $count++;
}
```

```
close(f);

$no_sent=$count;
#print "$maxClustNo  $maxDurNo  $maxnoCand  $maxClustLen $no_sent\
n";


while(($k,$v)=each %pair){
  if ($v >=10){
    print $k,"\t",$v,"\n";
  }
}


$iteration1=0;
$best_score=2;



while($best_score >= 2){
  $iteration++;
  undef %clusNo_bad;
  undef %clusNo_good;
  undef %word1_good;
  undef %word1_bad;
  undef %word2_good;
  undef %word2_bad;
  undef %pair_good;
  undef %pair_bad;
  undef %isEps1_good;
  undef %isEps1_bad;
  undef %isEps2_good;
  undef %isEps2_bad;
  undef %epspair_good;
  undef %epspair_bad;
  undef %dur1_good;
  undef %dur1_bad;
  undef %dur2_good;
  undef %dur2_bad;
  undef %Post1_good;
  undef %Post1_bad;
  undef %Post2_good;
  undef %Post2_bad;

  undef %diffPost_good;
  undef %diffPost_bad;

  undef %newCand_good;
undef %newCand_bad;

undef %Cand_good;
undef %Cand_bad;

undef %ClusterLen_good;
undef %ClusterLen_bad;

undef %score;

for($i=0; $i<$no_sent; $i++){

  $post1=int($sent[$i][7]*100);
  $post2=int($sent[$i][8]*100);
  $diffPost=int($sent[$i][11]*100);

  if($truth[$i] == $prediction[$i]){
    # the change is bad
    $clusNo_bad{$sent[$i][0]}{$truth[$i]}++;
    $word1_bad{$sent[$i][1]}{$truth[$i]}++;
    $word2_bad{$sent[$i][2]}{$truth[$i]}++;
    $pair_bad{$sent[$i][1]."_".$sent[$i][2]}{$truth[$i]}++;
    $isEps1_bad{$sent[$i][3]}{$truth[$i]}++;
    $isEps2_bad{$sent[$i][4]}{$truth[$i]}++;
    $epspair_bad{$sent[$i][3]."_".$sent[$i][4]}{$truth[$i]}++;

    $dur1_bad{$sent[$i][5]}{$truth[$i]}++;
    $dur2_bad{$sent[$i][6]}{$truth[$i]}++;

    $Post1_bad{$post1}{$truth[$i]}++;
    $Post2_bad{$post2}{$truth[$i]}++;

    $newCand_bad{$sent[$i][9]}{$truth[$i]}++;
    $Cand_bad{$sent[$i][10]}{$truth[$i]}++;
    $diffPost_bad{$diffPost}{$truth[$i]}++;
    $ClusterLen_bad{$sent[$i][12]}{$truth[$i]}++;
  }
  else{
    $change="$prediction[$i]:$truth[$i]";
    #print $change,"\n";
    $rule_type{$change}=1;  # 1->2  2->1
    $clusNo_good{$sent[$i][0]}{$change}++;
    $word1_good{$sent[$i][1]}{$change}++;
    $word2_good{$sent[$i][2]}{$change}++;
    $pair_good{$sent[$i][1]."_".$sent[$i][2]}{$change}++;
    $isEps1_good{$sent[$i][3]}{$change}++;
    $isEps2_good{$sent[$i][4]}{$change}++;
    $epspair_good{$sent[$i][3]."_".$sent[$i][4]}{$change}++;
```

```
   $dur1_good{$sent[$i][5]}{$change}++;
   $dur2_good{$sent[$i][6]}{$change}++;

   $Post1_good{$post1}{$change}++;
   $Post2_good{$post2}{$change}++;

   $newCand_good{$sent[$i][9]}{$change}++;
   $Cand_good{$sent[$i][10]}{$change}++;

   $diffPost_good{$diffPost}{$change}++;
   $ClusterLen_good{$sent[$i][12]}{$change}++;
  }
 }

 while(($change,$v)= each %rule_type){
  @b=split(/:/,$change);
  $pred=$b[0];

  print $change,"\n";

  $ruleEQ=$change.":0:eq"; # look at the first attribute and com
pare eq
with the value; do $k if true
  $ruleLS=$change.":0:ls";
  $ruleGT=$change.":0:gt";

  $rules{$ruleEQ}=1;
  $rules{$ruleLS}=1;
  $rules{$ruleGT}=1;


  for ($i=0; $i<=$maxClustNo; $i++){
   $score{$ruleEQ}{$i}= $clusNo_good{$i}{$change} -
$clusNo_bad{$i}{$pred};
   print $ruleEQ,"\t",$i,"\t", $clusNo_good{$i}{$change}," - ",

$clusNo_bad{$i}{$pred}," = $score{$ruleEQ}{$i}\n";
  }

  for ($i=1; $i<=$maxClustNo; $i++){
   for ($j=0; $j<$i; $j++){
    $score{$ruleLS}{$i}+=$score{$ruleEQ}{$j};
   }
   print $ruleLS,"\t",$i,"\t",$score{$ruleLS}{$i},"\n";
  }

  for ($i=$maxClustNo-1; $i>=0; $i--){
      for ($j=$i; $j<=$maxClustNo; $j++){
        $score{$ruleGT}{$i}+=$score{$ruleEQ}{$j};
      }
      print $ruleGT,"\t",$i,"\t",$score{$ruleGT}{$i},"\n";
    }

    $ruleEQ1=$change.":1:eq";
    $ruleEQ2=$change.":2:eq";
    $ruleEQ3=$change.":1_2:eq";

    $rules{$ruleEQ1}=1;
    $rules{$ruleEQ2}=1;
    $rules{$ruleEQ3}=1;


    while(($w1,$s1) = each %word1){
      #print $w1,"\t", $word1_good{$w1}{$change},"--\t--",
$word1_bad{$w1}{$pred},"\n";
      $score{$ruleEQ1}{$w1}= $word1_good{$w1}{$change} -
$word1_bad{$w1}{$pred};
    }

    while(($w1,$s1) = each %word2){
      #print $w1,"\t", $word2_good{$w1}{$change},"--\t--",
$word2_bad{$w1}{$pred},"\n";
      $score{$ruleEQ2}{$w1}= $word2_good{$w1}{$change} -
$word2_bad{$w1}{$pred};
    }

    while(($w1,$s1) = each %pair){
      #print $w1,"\t", $pair_good{$w1}{$change},"--\t--",
$pair_bad{$w1}{$pred},"\n";
      $score{$ruleEQ3}{$w1}= $pair_good{$w1}{$change} -
$pair_bad{$w1}{$pred};
    }

    $ruleEQ1=$change.":3:eq";
    $ruleEQ2=$change.":4:eq";
    $ruleEQ3=$change.":3_4:eq";

    $rules{$ruleEQ1}=1;
    $rules{$ruleEQ2}=1;
    $rules{$ruleEQ3}=1;

    $w1="yes";
    $w2="no";

    $score{$ruleEQ1}{$w1}= $isEps1_good{$w1}{$change} -
```

```
$isEps1_bad{$w1}{$pred};
    $score{$ruleEQ1}{$w2}= $isEps1_good{$w2}{$change} -
$isEps1_bad{$w2}{$pred};
    $score{$ruleEQ2}{$w1}= $isEps2_good{$w1}{$change} -
$isEps2_bad{$w1}{$pred};
    $score{$ruleEQ2}{$w2}= $isEps2_good{$w2}{$change} -
$isEps2_bad{$w2}{$pred};

    while(($w1,$s1) = each %epspair){
      $score{$ruleEQ3}{$w1}= $epspair_good{$w1}{$change} -
$epspair_bad{$w1}{$pred};
      print $w1,"\t", $epspair_good{$w1}{$change}," - ",
$epspair_bad{$w1}{$pred}," = $score{$ruleEQ3}{$w1}\n";
    }

    $ruleEQ=$change.":5:eq"; # look at the first attribute and com
pare eq
with the value; do $k if true
    $ruleLS=$change.":5:ls";
    $ruleGT=$change.":5:gt";

    $rules{$ruleEQ}=1;
    $rules{$ruleLS}=1;
    $rules{$ruleGT}=1;

    for ($i=0; $i<=$maxDurNo; $i++){
      $score{$ruleEQ}{$i}= $dur1_good{$i}{$change} -
$dur1_bad{$i}{$pred};
      print $ruleEQ,"\t",$i,"\t", $dur1_good{$i}{$change}," - ",
$dur1_bad{$i}{$pred}," = $score{$ruleEQ}{$i}\n";
    }

    for ($i=1; $i<=$maxDurNo; $i++){
      for ($j=0; $j<$i; $j++){
        $score{$ruleLS}{$i}+=$score{$ruleEQ}{$j};
      }
      print $ruleLS,"\t",$i,"\t",$score{$ruleLS}{$i},"\n";
    }

    for ($i=$maxDurNo-1; $i>=0; $i--){
      for ($j=$i; $j<=$maxDurNo; $j++){
        $score{$ruleGT}{$i}+=$score{$ruleEQ}{$j};
      }
      print $ruleGT,"\t",$i,"\t",$score{$ruleGT}{$i},"\n";
    }

    $ruleEQ=$change.":6:eq"; # look at the first attribute and com
pare eq
```

```
with the value; do $k if true
    $ruleLS=$change.":6:ls";
    $ruleGT=$change.":6:gt";

    $rules{$ruleEQ}=1;
    $rules{$ruleLS}=1;
    $rules{$ruleGT}=1;

    for ($i=0; $i<=$maxDurNo; $i++){
      $score{$ruleEQ}{$i}= $dur2_good{$i}{$change} -
$dur2_bad{$i}{$pred};
      print $ruleEQ,"\t",$i,"\t", $dur2_good{$i}{$change}," - ",
$dur2_bad{$i}{$pred}," = $score{$ruleEQ}{$i}\n";
    }

    for ($i=1; $i<=$maxDurNo; $i++){
      for ($j=0; $j<$i; $j++){
        $score{$ruleLS}{$i}+=$score{$ruleEQ}{$j};
      }
      print $ruleLS,"\t",$i,"\t",$score{$ruleLS}{$i},"\n";
    }

    for ($i=$maxDurNo-1; $i>=0; $i--){
      for ($j=$i; $j<=$maxDurNo; $j++){
        $score{$ruleGT}{$i}+=$score{$ruleEQ}{$j};
      }
      print $ruleGT,"\t",$i,"\t",$score{$ruleGT}{$i},"\n";
    }

    $ruleLS=$change.":7:ls";
    $ruleGT=$change.":7:gt";

    $rules{$ruleLS}=1;
    $rules{$ruleGT}=1;

    for ($i=0; $i<=100; $i++){
      for ($j=0; $j<$i; $j++){
        $score{$ruleLS}{$i}+=$Post1_good{$j}{$change} -
$Post1_bad{$j}{$pred};;
      }
      print $ruleLS,"\t",$i,"\t",$score{$ruleLS}{$i},"\n";
    }

    for ($i=99; $i>=0; $i--){
      for ($j=$i; $j<=100; $j++){
        $score{$ruleGT}{$i}+=$Post1_good{$j}{$change} -
$Post1_bad{$j}{$pred};;
      }
      print $ruleGT,"\t",$i,"\t",$score{$ruleGT}{$i},"\n";
    }

    $ruleLS=$change.":8:ls";
    $ruleGT=$change.":8:gt";

    $rules{$ruleLS}=1;
    $rules{$ruleGT}=1;

    for ($i=0; $i<=100; $i++){
      #print $ruleLS,"\t",$i,"\t", $Post2_good{$i}{$change},"--\t-
-",
$Post2_bad{$i}{$pred},"\n";
      for ($j=0; $j<$i; $j++){
        $score{$ruleLS}{$i}+=$Post2_good{$j}{$change} -
$Post2_bad{$j}{$pred};;
      }
      print $ruleLS,"\t",$i,"\t",$score{$ruleLS}{$i},"\n";
    }

    for ($i=99; $i>=0; $i--){
      for ($j=$i; $j<=100; $j++){
        $score{$ruleGT}{$i}+=$Post2_good{$j}{$change} -
$Post2_bad{$j}{$pred};;
      }
      print $ruleGT,"\t",$i,"\t",$score{$ruleGT}{$i},"\n";
    }

    $ruleEQ=$change.":9:eq";
    $rules{$ruleEQ}=1;

    for ($i=2; $i<=3; $i++){
      print $ruleEQ,"\t",$i,"\t", $newCand_good{$i}{$change},"--\t
--",
$newCand_bad{$i}{$pred},"\n";
      $score{$ruleEQ}{$i}= $newCand_good{$i}{$change} -
$newCand_bad{$i}{$pred};
    }

    $ruleEQ=$change.":10:eq";
    $ruleLS=$change.":10:ls";
    $ruleGT=$change.":10:gt";

    $rules{$ruleEQ}=1;
    $rules{$ruleLS}=1;
    $rules{$ruleGT}=1;

    for ($i=0; $i<=$maxCandNo; $i++){
```

```
   $score{$ruleEQ}{$i}= $Cand_good{$i}{$change} -
$Cand_bad{$i}{$pred};
   print $ruleEQ,"\t",$i,"\t", $Cand_good{$i}{$change}," -",
$Cand_bad{$i}{$pred}," = $score{$ruleEQ}{$i}\n";
  }

  for ($i=1; $i<=$maxCandNo; $i++){
   for ($j=0; $j<$i; $j++){
    $score{$ruleLS}{$i}+=$score{$ruleEQ}{$j};
   }
   print $ruleLS,"\t",$i,"\t",$score{$ruleLS}{$i},"\n";
  }

  for ($i=$maxCandNo-1; $i>=0; $i--){
   for ($j=$i; $j<=$maxCandNo; $j++){
    $score{$ruleGT}{$i}+=$score{$ruleEQ}{$j};
   }
   print $ruleGT,"\t",$i,"\t",$score{$ruleGT}{$i},"\n";
  }

  $ruleLS=$change.":11:ls";
  $ruleGT=$change.":11:gt";

  $rules{$ruleLS}=1;
  $rules{$ruleGT}=1;

  for ($i=0; $i<=100; $i++){
   for ($j=0; $j<$i; $j++){
    $score{$ruleLS}{$i}+=$diffPost_good{$j}{$change} -
$diffPost_bad{$j}{$pred};;
   }
   print $ruleLS,"\t",$i,"\t",$score{$ruleLS}{$i},"\n";
  }

  for ($i=99; $i>=0; $i--){
   for ($j=$i; $j<=100; $j++){
    $score{$ruleGT}{$i}+=$diffPost_good{$j}{$change} -
$diffPost_bad{$j}{$pred};;
   }
   print $ruleGT,"\t",$i,"\t",$score{$ruleGT}{$i},"\n";
  }


  $ruleEQ=$change.":12:eq";
  $ruleLS=$change.":12:ls";
  $ruleGT=$change.":12:gt";

  $rules{$ruleEQ}=1;
```

```
    $rules{$ruleLS}=1;
    $rules{$ruleGT}=1;


    for ($i=0; $i<=$maxClustLen; $i++){
      $score{$ruleEQ}{$i}= $ClusterLen_good{$i}{$change} -
$ClusterLen_bad{$i}{$pred};
      print $ruleEQ,"\t",$i,"\t", $ClusterLen_good{$i}{$change},"
- ",
$ClusterLen_bad{$i}{$pred}," = $score{$ruleEQ}{$i}\n";
    }

    for ($i=1; $i<=$maxClustLen; $i++){
      for ($j=0; $j<$i; $j++){
        $score{$ruleLS}{$i}+=$score{$ruleEQ}{$j};
      }
      print $ruleLS,"\t",$i,"\t",$score{$ruleLS}{$i},"\n";
    }

    for ($i=$maxClustLen-1; $i>=0; $i--){
      for ($j=$i; $j<=$maxClustLen; $j++){
        $score{$ruleGT}{$i}+=$score{$ruleEQ}{$j};
      }
      print $ruleGT,"\t",$i,"\t",$score{$ruleGT}{$i},"\n";
    }

    $ruleEQ=$change.":5_6_11:eq";

  }

  print "OUT OF HERE\n";

  $best_rule="";
  $best_score=1;

  while(($k,$v)=each %rules){
    *a=$score{$k};

    while(($k1,$v1)=each %a){
      if ($v1 >= $best_score){
        print $k,"\t",$k1,"\t",$v1,"\n";
        $best_score = $v1;
        $best_rule=$k.":".$k1;
      }
    }
  }

  if ($best_score >=1 && $best_rule ne ""){
```

```
  print "ITERATION: $iteration\n";
  print "BEST SCORE $best_score\n";
  print "BEST RULE $best_rule\n";

  print fout "$best_rule  $best_score\n";


  @c=split(/:/,$best_rule);
  $attribute=$c[2];
  $source=$c[0];
  $target=$c[1];
  $value=$c[4];
  if ($attribute eq "7" || $attribute eq "8" || $attribute eq "1
1"){
   $value=$value/100;
  }
  $comparison=$c[3];


  for ($i=0; $i< $no_sent; $i++){
   if ($comparison eq "eq"){
    if ($attribute=~/_/){
     if ($value=~/_/){
      @comp=split(/_/,$attribute);
      @val=split(/_/,$value);
      if ($sent[$i][$comp[0]] eq $val[0] && $sent[$i][$comp[
1]] eq
$val[1] && $prediction[$i] eq $source){
       $prediction[$i] = $target;
       print "I changed $source to $target for $i\n";
      }
     }
     else{
      print "ERROR $attribute $value\n";
      last;
     }
    }
    else{
     if ($sent[$i][$attribute] eq $value && $prediction[$i] e
q
$source){
      $prediction[$i] = $target;
      print "I changed $source to $target for $i\n";
     }
    }
   }
```

```
      elsif($comparison eq "ls"){
        if ($sent[$i][$attribute] <= $value && $prediction[$i] eq
$source){
          $prediction[$i] = $target;
          print "I changed $source to $target for $i\n";
        }
      }
      elsif($comparison eq "gt"){
        if ($sent[$i][$attribute] >= $value && $prediction[$i] eq
$source){
          $prediction[$i] = $target;
          print "I changed $source to $target for $i\n";
        }
      }
      else{
        print "ERROR: unknown comparison $comparison! \n";
        exit;
      }
    }
  }
  if ($best_rule eq ""){
    $best_score=-1;
  }
}
```